

Sampling Urban Mobility through On-line Repositories of GPS Tracks

Michał Piórkowski

LCA4, School of Computer and Communication Sciences
EPFL, 1015 Switzerland
michal.piorkowski@gmail.com

ABSTRACT

We analyze urban mobility by relying on the short-term mobility traces gathered from a publicly available web-based repository of GPS tracks - the Nokia Sports Tracker service¹. All mobility traces are obtained from a set of `km1` files². We show how the data collected voluntarily by individuals, equipped with GPS-enabled mobile phones, can be used to infer accurate, large-scale footprint of urban mobility. This method, unlike others - for example, personal interviewing, is more scalable and less time consuming. It exploits the fact that the on-line masses are willing to share their experience with others.

We present a set of heuristics used to filter out bogus tracks from the dataset. We show that the mobility patterns, inferred from the remaining, credible, short-term mobility traces have macroscopic characteristics similar to the characteristics of mobility patterns retrieved from the long-term mobility traces, gathered in different urban environments. The results of our analysis lead to a proposal for creating city-specific mobility profiles. We discuss how such profiles could help improve location privacy and help develop new context-aware applications and services for mobile users.

Categories and Subject Descriptors

G.3 [Probability and Statistics]: Nonparametric statistics; H.3.3 [Information Storage and Retrieval]: [Information Filtering, Clustering]

General Terms

Measurement, Modeling

Keywords

Mobility Modeling, Privacy, Context-Awareness

¹<http://sportstracker.nokia.com>

²<http://www.opengeospatial.org/standards/kml>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

ACM HotPlanet '09 Krakow, Poland

Copyright 2009 ACM 978-1-60558-689-2/09/06 ...\$10.00.

1. INTRODUCTION

In the year 2008, Nokia planned to ship out 35 million GPS-enabled devices³. The proliferation of such devices, together with the implications beyond the hype of the iPhone, allows many location-based services to flourish⁴. Nokia Sports Tracker is one of these services: mobile phone owners use it to upload their activity traces to a web-based GPS track repository to store and share with others.

In this paper we show that from such on-line repositories it is possible to retrieve short-term mobility traces and to infer mobility pattern specific to a certain region. We present the techniques we developed and give a general overview of how to collect and postprocess GPS tracks from a publicly available repository.

The datasets with long-term mobility traces contain locations of individuals recorded on a regular basis, independent on their activities. This is in contrast with our dataset, containing short-term mobility traces. We observe that the access to datasets with short-term mobility traces is facilitated by developments in Web 2.0 technologies. This simplifies mobility modeling in urban scenario, where large scale datasets are required for meaningful analysis. We observe that although such short-term mobility traces contain only a snapshot of people's mobility, they can be used to infer collective mobility patterns that unveil macroscopic characteristics similar to characteristics of mobility patterns derived from datasets containing long-term mobility.

We believe that studying mobility patterns by using the publicly available short-term individual mobility traces should not raise significant privacy concerns as users of services such as the Nokia Sports Tracker accept to publish information about their trips⁵. Note that this is not the case for the long-term mobility traces, gathered without an explicit consent from the end users [4]. The primary deterrents to collection of large-scale mobility data, i.e., user's privacy and recruitment, are automatically circumvented when the data is collected on a short-term basis through a service fully controlled by each user.

Many large-scale mobility traces are coarse-grained both in time and location. For example the most fine-grained

³Olli-Pekka Kallasvuo, President and CEO, Nokia
Nokia Annual General Meeting, 8 May 2008

⁴The number of location-aware applications added to the iPhone App Store increased from 500 to 2500 between October 2008 and May 2009:
<http://www.skyhookwireless.com/locationapps>

⁵End User Software Agreement for Nokia Sports Tracker:
<http://sportstracker.nokia.com/nts/eusa.do>

dataset studied by Gonzales et al. in [6] contains traces of 206 mobile phone users whose locations, represented by the coordinates of the closest base station, are recorded every 2 hours for a period of 6 months. In contrast the traces shared on-line through the Nokia Sports Tracker service are fine-grained in both time (location update time interval in seconds) and space (user geo-location instead of base station coordinates).

It is worth noticing that very often the large-scale mobility traces contain only location information. The dataset available through the Nokia Sports Tracker service contains also contextual information. Users explicitly specify their activities associated with each route, e.g. walking, running, cycling, driving, etc. This allows us to associate the known context of people's activities with certain locations. It is crucial for many context-aware services, which are sensitive to context much more than to location. Using the Nokia Sports Tracker dataset we show how certain activities cluster in specific areas of the city. The results of our study lead, inter alia, to a proposal for improving the location privacy in mobile ad hoc networks operating in urban environment.

To summarize, we make the following contributions:

- We present the guidelines for the retrieval of urban mobility patterns from a web-based repository of GPS tracks.
- We show that mobility patterns inferred from short-term mobility traces share macroscopic characteristics similar to characteristics of mobility patterns extracted from long-term traces.
- We analyze the mobility traces from three European capitals to show the spatial clustering of activities.
- We introduce the concept of a *city-specific mobility profile*.
- We discuss how such a mobility profile can be used to design context-aware applications and how it can help achieve significant location privacy in mobile networks.

The paper is organized as follows. In Section 2 we discuss the related work. In Section 3 we present the dataset we retrieve plus the heuristics we use for filtering. Next, in Section 4 we show the evidence that both short-term and long-term mobility traces unveil similar macroscopic mobility characteristics. We also introduce so-called *city-specific mobility profile* as a way to capture urban mobility footprint. In Section 5 we discuss the potential applications of the city-specific mobility profile inferred from a web-based repository of GPS tracks. Finally in Section 6 we conclude our work and we present possible future directions for this paper.

2. RELATED WORK

There exist many publicly available datasets of human mobility, e.g. Hagggle iMote [14] or Reality Mining [15]. The main problem with these datasets is that they are too small to study large-scale mobility. This is because in order to gather such traces many individuals have to be recruited and equipped with dedicated devices that record their location and (sometimes) contextual information. In this work we propose to focus on publicly available repositories of GPS

tracks as a resource for studying human mobility in urban environments.

In fact there is a plethora of on-line repositories, similar to Nokia Sports Tracker, e.g. [19, 20, 21]. To the best of our knowledge, this paper is the first to study city-wide human mobility based solely on the short-term mobility traces obtained from an online repository for GPS-based tracks.

There have been many successful attempts to study the mobility of many individuals in an urban environment. Balazinska et al. in [1] present results from a four-week trace collected in a large corporate environment. They use the 802.11b local-area network to track user associations with 177 access points. They are able to infer long-term mobility patterns of 1366 unique MAC addresses together with detailed information on the amount of data transferred. Note that the dataset available through Nokia Sport Tracker does not allow for the retrieval of the long-term mobility traces. However it provides more accurate location information. In addition, user activity-related context is also available. Gonzalez et al. in [6] study mobility patterns of 100'000 individuals using the data set disclosed by a telecom operator. This raised many privacy concerns as mobility data was provided without an explicit consent from the end users [4]. Note that in the case of the Nokia Sports Tracker the privacy issues are relaxed by users. In [12], Reades et al. study mobility patterns by using a dataset also provided by a telecom company. This dataset consists of aggregated and anonymized mobile phone use, more specifically the network bandwidth use collected at the antenna level. In contrast to the dataset used by Reades et al., the Nokia Sports Tracker dataset allows us to infer city-wide mobility patterns at a more fine-grained level.

A similar analysis of collective mobility patterns in urban environment was done by Kaltenbrunner et al. in [8]. Authors obtain the data by periodically mining **km1** files accessible through the Bicing website [18]. They analyze urban mobility by tracking bicycles at the stations of the community bicycle program in Barcelona. Using dedicated heuristics, they are able to detect temporal patterns in mobility as well as identify residential, business and leisure areas. The data provided by the Nokia Sports Tracker service allows us to characterize mobility not only at the origin and destination of a trip, but also in between.

3. NOKIA SPORTS TRACKER

Nokia Sports Tracker is a location-based service provided by Nokia, which allows users to track their routes, speed and timings while engaging in various activities. Each user equipped with a GPS-enabled mobile phone has to install a dedicated piece of software that allows for the recording of the track and for uploading it to the on-line repository. Each publicly available track is a compressed version of the original track, downloadable as a **km1** file. In Table 1 we show what information is stored in each **km1** file. Each route is represented as a collection of waypoints where each waypoint is a three-tuple: (longitude, latitude, altitude). Users can associate with each route a specific activity, e.g. walking, running, cycling, driving, etc.

3.1 Dataset Retrieval

For the purpose of this study we collected approximately 125'000 tracks recorded between July 2007 and September 2008. This large dataset of **km1** files contains traces from

Table 1: Information stored in each **kml** file from Nokia Sports Tracker online repository

user	nickname
type of activity	e.g. walking
start time	date hh:mm
duration	hh:mm:ss
traveled distance	km
average speed	km/h
maximum speed	km/h
average pace	1/km
maximum pace	1/km
the route	set of waypoints

all over the world. As we focus on analyzing urban mobility, we preselect only the traces recorded in three European capitals: Helsinki (8330 traces), Stockholm (1976 traces), and London (2918 traces). The resulting dataset combines traces from different activities (running, walking, cycling, etc.), and as such only provides occasional snapshots of a person’s long-term mobility.

3.2 Postprocessing

The purpose of postprocessing is to remove tracks that are not credible for a meaningful analysis. We are aware that GPS-enabled devices operating in an urban environment, where the sky is often occluded, might generate tracks with many errors. To the best of our knowledge, the Nokia Sports Tracker service does not currently provide an automatic correction of such errors in the uploaded GPS tracks. Thus the resulting **kml** files may be bogus. Here we present the set of heuristics used to filter out the inappropriate tracks.

In order to be able to calculate distances between waypoints for each track we convert geo-coordinates to Cartesian coordinates using Vincenty’s formulae [16]. This allows us to compare if the distance as given in the **kml** file is similar to the one calculated directly from the route. Also, it is important to verify if the average speed from the **kml** file is comparable to the speed calculated directly from the trace. Unfortunately the **kml** files do not contain information about the time difference between two consecutive waypoints (cf. Table 1). However, knowing that the default period for logging geo-coordinates by a Nokia GPS-enabled device is one second and knowing that on average each trace is compressed at the server side to a trace with up to approximately 200 waypoints only, we are able to find the time difference between consecutive waypoints. Note that we have reverse engineered the latter by studying the characteristics of routes from the collected **kml** files.

Our first filter is designed to remove from the raw dataset all tracks that are short in time (less than 1 minute) and in distance (less than 500 meters), because they may not be credible due to the insufficient amount of ephemeris data required to correctly calculate the geo-location. Next, we omit all the traces for which the maximum speed is lower than the average speed. We apply this filter twice - for the first time on the speeds provided in the corresponding **kml** file and for the second time on the speeds calculated from the track. Next, we verify the difference between the total traveled distance from the **kml** file and the distance calculated from the route relative to the latter. If it is larger than 30%, then the corresponding track is omitted. Finally,

Table 2: Filtering spurious tracks

		Raw [km/h]		Filtered [km/h]	
		$\hat{\mu}_v$	$\hat{\sigma}_v^2$	$\hat{\mu}_v$	$\hat{\sigma}_v^2$
Helsinki	Walking	7.75	180.15	6.37	46.11
	Running	10.83	551.24	9.85	4.10
	Cycling	18.81	69.82	17.59	35.26
Stockholm	Walking	6.65	191.22	6.18	10.52
	Running	11.19	45.59	10.35	11.89
	Cycling	18.16	41.70	16.98	23.82
London	Walking	8.08	321.23	7.48	24.70
	Running	11.05	11.70	11.07	6.33
	Cycling	19.27	614.94	18.29	36.62

we also apply the same procedure for the average speed.

As a result, we obtain collection of *clean* traces: Helsinki - 3757 traces recorded by 731 users, Stockholm - 1056 traces recorded by 157 users and London - 2488 traces recorded by 504 users. To verify whether the resulting dataset contains credible GPS tracks, we perform the following test. We compare the population-based mean $\hat{\mu}_v$ and the variance $\hat{\sigma}_v^2$ of the average speed v before and after filtering the dataset separately for each city and in addition for each activity (cf. Table 2). We focus on the three most characteristic and popular activities for the urban environment, i.e., walking, running and cycling. We omit other activities such as driving mainly because the size of the corresponding collection of tracks is too small for meaningful analysis. Also other activities such as orienteering or rowing are not very specific for the urban environment thus they are also omitted. The typical average speed for walking varies between 4.7 and 8.3 km/h [3], for running - between 9 and 15.2 km/h [3] and for cycling - between 15 and 25 km/h [17]. By comparing the $\hat{\mu}_v$ and $\hat{\sigma}_v^2$ calculated for the raw dataset with those calculated for the filtered one we are able to see how effective our filtering heuristics are. The significant reduction in the variability of the average speed (cf. Table 2 column 4 vs. 6) suggests that our heuristics work well.

Note that instead of removing the spurious tracks we could improve the track accuracy by reconciling a user’s location with the underlying map by using map matching methods. This is however out of scope of this paper. The reader is kindly referred to the rich body of related literature, e.g. [11].

4. SHORT-TERM URBAN MOBILITY

Here we verify if a dataset from the Nokia Sports Tracker service unveils mobility characteristics similar to other datasets containing large-scale and long-term mobility traces. Specifically, we compare mobility patterns derived from the Nokia Sports Tracker dataset with the two mobility traces we obtained from two large taxi companies from Warsaw [13] and San Francisco [9] - they both contain GPS-based tracks of taxi fleet (>500 cabs) recorded for a period of more than a month. For more details about the traces we refer the reader to the original papers. We are specifically interested in the node distribution and its impact on connectivity in mobile networks operating in an urban environment. It has been shown that the node spatial distribution is heterogeneous and thus can lead to network partitioning [13].

4.1 Spatial Distribution of Nodes

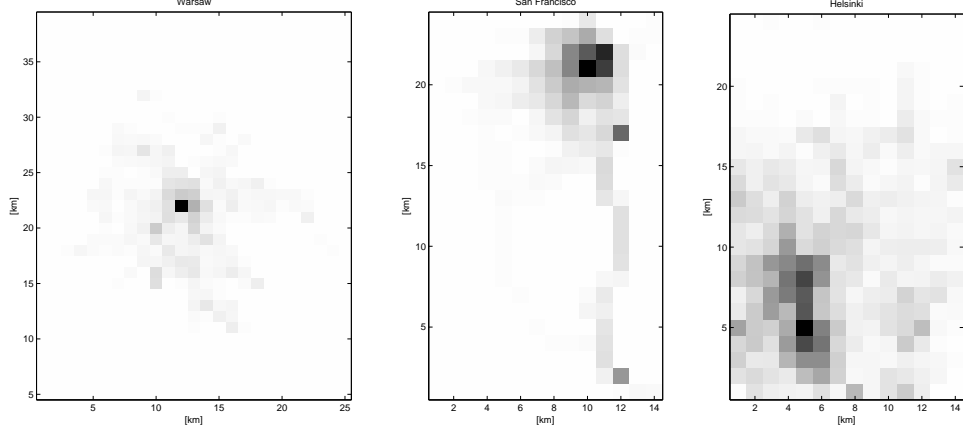


Figure 1: Density maps: spatial histogram of cell counts for Warsaw, San Francisco and Helsinki. A darker pixel indicates higher count.

Here we show that the spatial distribution of nodes in an urban environment is heavy tailed. For each dataset, we look at the distribution of counts of nodes visits to a grid of square cells of equal size. For this analysis, we first divide the whole region of Warsaw, San Francisco and Helsinki into a set of square cells of equal size (1000×1000 m). We focus specifically on Helsinki because the dataset for this city is the largest among those we extracted from the Nokia Sports Tracker repository, thus it is statistically more significant. Call C_z - the number of waypoints that fall into cell z - a count for cell z . The set of cells with their corresponding counts C_z can be interpreted as a *density map*. In Figure 1 we present density maps for Warsaw, San Francisco and Helsinki. For each cell z we find $c(z) = \frac{C_z}{\sum C_z}$, which can be interpreted as an empirical probability that a random waypoint falls into the cell z . Figure 2 shows the empirical complementary cumulative distribution function (CCDF) of $c(z)$ for all the three datasets. Linear behavior on the plot is evidence for a heavy tail distribution, which implies that some cells have a population density much above the average. We find that the tail index for Warsaw is $\lambda_W = 1.65$, for San Francisco - $\lambda_{SF} = 1.45$ and for Helsinki - $\lambda_H = 1.3$.

4.2 Activity Mapping

Recall that the Nokia Sports Tracker software installed on a mobile phone allows each user to associate a specific activity with each track. Here we focus again on the three most characteristic activities for the urban environment: walking, running and cycling. In Figure 3 we present the density maps for each activity and for each city separately. Visual inspection allows us to draw a conclusion that each activity has its own spatial pattern.

4.3 Network Connectivity

Unlike the datasets with explicit connectivity information [14, 15], all three datasets under our investigation do not contain any information about node connectivity. Hence, we need to infer the contacts between nodes if we want to study the connectivity.

Let us first define the *connectivity graph*. The mobile nodes and the corresponding wireless links define the *connectivity graph* $G(V, E)$, where $V(G)$ is the set of mobile

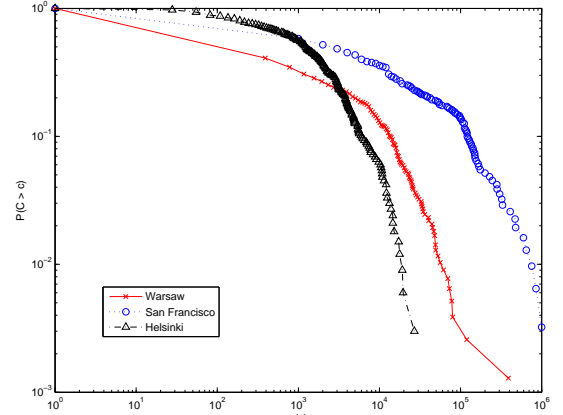


Figure 2: Empirical CCDF of $c(z)$ in log-log scale for three cities Warsaw, San Francisco and Helsinki.

nodes and $E(G)$ is the set of radio links between mobile nodes, i.e., $E(G) = \{e = (i, j) | d_{ij} \leq r\}$. We assume that r is fixed and represents the maximal radius allowed by power constraints. We define H_k as a component of G , with $\mathbb{C}(G) = \{H_1, H_2, \dots, H_K\}$ the set of all components, i.e., $G = \bigcup_{k=1}^{K=K} H_k$.

To show the emergence of connectivity islands in urban environments, we study how the structure of the connectivity graph changes with increasing r . More precisely, we focus on the size of components $|H_k|$ present in the connectivity graph $G(V, E)$ generated at a random time instant from all three data sets. For this experiment, we *densify* our trace in the following way. We assume that the sampled nodes are a representative subset of the overall traffic, and that their movements are stationary and ergodic. Under this assumption, we can generate a denser sample of node locations (in our case, 3000 nodes) by sampling uniformly at random from the entire dataset. This denser sample brings out more detail in the connectivity graph.

In Figure 4 we show the results. We observe a clear trend in all the cases - the number of connected components of a size larger than 10 rises quickly with the communication

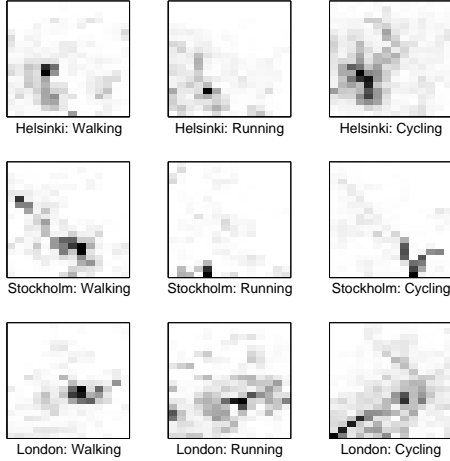


Figure 3: Visualization of a city-specific mobility profile for Helsinki, Stockholm and London - density maps specific for different activities. A darker pixel indicates higher count.

range r towards a maximum; past this critical value, the number of components starts to decrease, because smaller components start coalescing into larger ones. However, this decrease tends to be slow; even when r becomes a multiple of the critical value, there are still many components left. This is because node locations are not distributed uniformly in space, which prevents percolation into a single giant cluster when r grows. Thus, we conclude that in an urban environment disconnected network topologies, with a large number of components, seem to be a robust phenomenon that persists over a wide range of radio ranges.

4.4 Discussion

Short-term mobility traces, produced by a large group of Nokia Sports Tracker users, unveil macroscopic characteristics of collective urban mobility similar to the long-term mobility traces generated by taxi fleets. We conjecture that it can be attributed to features of the urban environment, where urban landscapes shape the way people move. A similar observation was made by Jiang et al. in [7], that it is the underlying road network topology that governs the scaling behavior of the node spatial distribution. Interestingly, despite differences in the mobility type and city topologies, the connectivity characteristics of city-wide mobile ad hoc networks seem to be very similar.

4.5 City-Specific Mobility Profile

Our analysis shows that the short-term mobility traces unveil similar macroscopic mobility characteristics as the long-term traces. Hence, they appear to be useful for inferring urban mobility patterns. Therefore, we postulate that the publicly available repository of GPS tracks can be used to create so-called *city-specific mobility profiles* that represent a collection of places to which context and mobility specific attributes and corresponding values are assigned (cf. Figure 3). A city-specific mobility profile can be used, for example, to produce synthetic mobility traces, characteristic for a certain region. For example, in the case of the Heterogeneous Random Walk mobility model [10], this can be done by inferring mobility model parameters from the mobility trace,

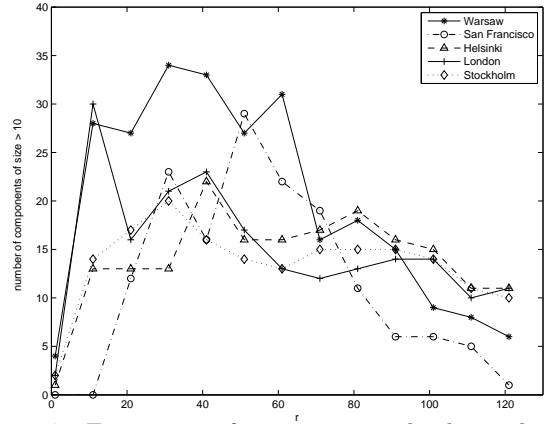


Figure 4: Emergence of connectivity islands in the connectivity graph $G(V, E)$ for different values of r [m].

specifically the location and size of large connected components. Additionally, a city-specific mobility profile can help enhance the location privacy in mobile networks and also design context-aware applications. We discuss these in more detail in the next section.

5. POTENTIAL APPLICATIONS

5.1 Improving Location Privacy

In mobile ad hoc networks, third parties are capable of tracking a mobile node's location by monitoring the pseudonyms used for identification. An existing solution to protect the location privacy of mobile nodes suggests changing pseudonyms in regions called *Mix Zones* [2]. Each Mix Zone is an anonymizing region that obfuscates the relation between entering and exiting nodes. It has been shown that the location privacy increases with two factors: (i) the number of nodes in a Mix Zone and (ii) the similarity of a mobility profile of individuals inside a Mix Zone. It means that the location privacy achieved in mobile network highly depends on the placement of the Mix Zones. Recently, Freudiger et al. have shown in [5] that the city-specific mobility profile could be used to find the optimal placement of Mix Zones so that the location privacy is highly improved, i.e., the tracking success by an adversary is decreased when Mix Zones are well placed.

Nokia Sports Tracker users represent a specific part of the population ready to relax their privacy in order to have the opportunity to share with others information about their past trajectories. Their generosity can be exploited by application developers to design better location privacy schemes, e.g. by identifying *hot-spot* locations specific to a certain activity (cf. Figure 3). In other words mobile users concerned about their location privacy could rely on a Mix Zone infrastructure designed by exploiting the mobility profiles of Nokia Sports Tracker users.

5.2 Designing Context-Aware Services

In many context-aware services it is not enough to rely solely on the location information. This is because there is more that defines an individual, his preferences and state, than his location. We believe that city-specific mobility profiles should help developers and service providers to build

better context-aware applications and services.

We know that context information is more difficult to obtain than a user's location. Fortunately Nokia Sports Tracker users, apart from relaxing their location privacy, volunteer to unveil their context. Hence mining the mobility traces of Nokia Sports Tracker users, allows us to leverage both - their location and context. For example, vehicle traffic updates could be context-aware, e.g. only while driving or cycling users would be provided imperceptibly with such updates.

6. CONCLUSIONS AND FUTURE WORK

We propose a set of guidelines to collect and filter short-term mobility traces from publicly available repositories of GPS tracks. To the best of our knowledge, this is the first work that exploits such repositories to study large-scale urban mobility. Our method relies on the fact that mobile users, subscribed to an on-line service dedicated to location tracking, agree to relax their privacy to be able to share their tracks with others. This data collection approach complements traditional techniques of gathering mobility traces. We have shown that the collection of sporadic mobility traces of many users unveils characteristics of mobility patterns similar to long-term mobility traces. Finally we propose an abstraction called city-specific mobility profile, which can be used to enhance location privacy and to design better context-aware services.

We believe that building city-specific mobility profiles based on the short-term mobility traces is a good starting point not only to find optimal locations of Mix Zones but also to design better context-aware services. However, more data needs to be collected for this purpose. We intend to gather more data also for the purpose of analyzing the activity cycles specific to particular locations. We also intend to focus more on the individual mobility - analyzing this information would allow us to build new personalized context-aware services. We would also like to revise our assumption about the representativeness of the city-specific mobility profile inferred from the web services like Nokia's. Although the Nokia Sports Tracker is aimed at sports users, we believe that the number of generic users of this service is increasing. Walking or cycling is not only associated with recreation or training. In fact, the most common type of cycling is utility cycling. This suggests that with time the Nokia Sports Tracker user base may become a representative sample of the city population.

Acknowledgments

The authors would like to thank Maciej Kurant, Marcin Poturalski, Julien Freudiger and Holly Cogliati for their help with Python and for help in improving the manuscript. Also, we would like to thank anonymous reviewers who provided us with very helpful comments. Finally, we are indebted to the Nokia Sports Tracker developers as well as to the members of the Nokia Sports Tracker community that share their experience on-line.

The work presented in this paper has been supported (in part) by the National Competence Center in Research on Mobile Information and Communication Systems (NCCR-MICS), a center supported by the Swiss National Science Foundation under grant number 5006-67322.

7. REFERENCES

- [1] M. Balazinska and P. Castro. Characterizing Mobility and Network Usage in a Corporate Wireless Local-Area Network. In *ACM MobiSys*, 2003.
- [2] A. Beresford and F. Stajano. Mix-Zones: User Privacy in Location-aware Services. In *Proceedings of PerSec*, 2004.
- [3] D. M. Bramble and D. E. Lieberman. Endurance running and the evolution of *Homo*. *Nature*, 432:345–352, November 2004.
- [4] J. E. Dobson. Big Brother has evolved. *Nature*, 458:968, April 2009.
- [5] J. Freudiger, R. Shokri, and J.-P. Hubaux. On the Optimal Placement of Mix Zones. In *The 9th Privacy Enhancing Technologies Symposium*, 2009.
- [6] M. Gonzalez, C. Hidalgo, and A.-L. Barabasi. Understanding Individual Human Mobility Patterns. *Nature*, 453:779–782, June 2008.
- [7] B. Jiang, J. Yin, and S. Zhao. Characterizing Human Mobility Patterns over a Large Street Network. <http://arxiv.org/abs/0809.5001>, September 2008.
- [8] A. Kaltenbrunner, R. Meza, J. Grivolla, J. Codina, and R. Banchs. Bicycle cycles and mobility patterns - Exploring and characterizing data from a community bicycle program. <http://arxiv.org/abs/0810.4187v1>, October 2008.
- [9] M. Piórkowski, N. Sarafijanovic-Djukic, and M. Grossglauser. CRAWDAD data set epfl/mobility (v. 2009-02-24). <http://crawdad.cs.dartmouth.edu/epfl/mobility>.
- [10] M. Piórkowski, N. Sarafijanovic-Djukic, and M. Grossglauser. A Parsimonious Model of Mobile Partitioned Networks with Clustering. In *THE 1st International Conference on COMMUNICATION SYSTEMS and NETWORKS*, January 2009.
- [11] M. A. Quddus, W. Y. Ochieng, L. Zhao, and R. B. Noland. A general map matching algorithm for transport telematics applications. *GPS Solutions*, 7(3):157–167, December 2003.
- [12] J. Reades, F. Calabrese, A. Sevtsuk, and C. Ratti. Cellular census: Explorations in urban data collection. *Pervasive Computing*, 6(3):30–38, 2007.
- [13] N. Sarafijanovic-Djukic, M. Piórkowski, and M. Grossglauser. Island Hopping: Efficient Mobility-Assisted Forwarding in Partitioned Networks. In *IEEE SECON'06*, volume 1, pages 226–235, 2006.
- [14] J. Scott, R. Gass, J. Crowcroft, P. Hui, C. Diot, and A. Chaintreau. CRAWDAD trace set cambridge/haggle/imote (v. 2006-09-15). <http://crawdad.cs.dartmouth.edu/cambridge/haggle/imote>.
- [15] The Reality Mining Dataset. <http://reality.media.mit.edu/dataset.php>.
- [16] T. Vincenty. Direct and Inverse Solutions of Geodesics on the Ellipsoid with Application of Nested Equations. *Survey Review*, 23(176):88–93, April 1975.
- [17] D. G. Wilson and J. Papadopoulos. *Bicycling Science*. The MIT Press, 3 edition, 2004.
- [18] Bicing. <http://www.bicing.com>.
- [19] GPSed. <http://gpsed.com>.
- [20] MapMyRide. <http://www.mapmyride.com>.
- [21] Trailguru. <http://www.trailguru.com>.